



DATA SCIENCE MEETS RETAIL: PREDICTING BIG MARTS'S SALES

*Shaik Mulla Almas¹, G. Jayanth Satya², D. Pavan Kumar³, M. Yashwanth⁴, N. Sai Subhash⁵, Mohammed Asrar⁶

¹ Assistant Professor, Department of IT, Vasireddy Venkatadri Institute of Technology, Guntur, India (Corresponding Author)

^{2,3,4,5,6} B.Tech. Students, Department of IT, Vasireddy Venkatadri Institute of Technology, Guntur, India

ABSTRACT

In the modern era, shopping malls and Big Marts maintain a comprehensive record of their sales data for each individual item. This data is utilized to predict future customer demand and update inventory management. These records are stored in data warehouses, containing a vast amount of customer data and item attributes. By mining this data, anomalies and frequent patterns can be identified. The resulting information can be leveraged to forecast future sales volume using various machine learning techniques, specifically for retailers like Big Mart. In this study, we propose a predictive model that utilizes the XG boost Regressor technique to forecast sales for companies similar to Big Mart. It one the affordable method when comparative to other methods Our findings indicate that this model outperforms existing models in terms of performance.

KEYWORDS: XG Boost, Big Mart Sales, Data Warehouses.

INTRODUCTION

The Big Mart, a prominent supermarket chain with nationwide presence, has issued a challenge to all Data Scientists. The challenge entails developing a model capable of accurately predicting the sales of each product in every store. In order to accomplish this, Big Mart has obtained sales data from Kaggle, encompassing a wide range of products across multiple stores in different cities. By leveraging this valuable information, the corporation aims to identify the pivotal products and stores that significantly impact their sales. This knowledge will enable them to implement appropriate strategies to ensure the success of their business. Notably, the current machine learning algorithm is highly sophisticated, offering a plethora of methods to forecast sales for organizations of any nature. This advanced algorithm proves particularly advantageous in mitigating the challenges associated with low-cost prediction techniques..

MATERIALS AND METHODS

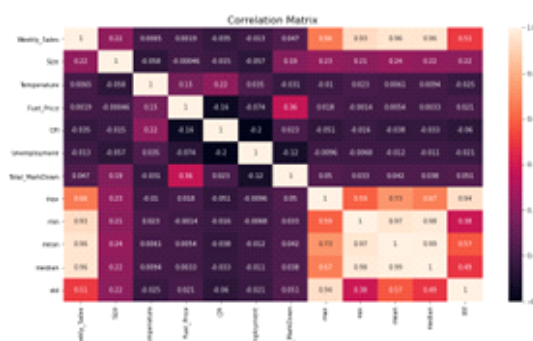
The methods used for sales prediction with XGBoost include data preprocessing, splitting the data into training and testing sets, training an XGBoost regression model, and evaluating the model's performance. Data preprocessing involves steps like data cleaning, feature engineering, data scaling and transformation, and encoding categorical variables. Splitting the data into training and testing sets allows for training the model on the training set and evaluating its performance on the testing set. The XGBoost model, known for its high performance in regression and classification tasks, is trained using the training data. Finally, the model's performance is evaluated using appropriate metrics such as Mean Absolute Error.

RESULTS

The Results section helps us to for better analysis and prediction of the sales and enhances the business sales and increases market share in the public. The potential benefits of utilizing XGBoost for sales prediction in Big Mart are emphasized, with a particular focus on its advantages. However, it is essential to underscore the importance of conducting further validation.



The co-relation matrix



DISCUSSION

Discussion Sales prediction using XGBoost is a valuable application for a retail giant like Big Mart. XGBoost is a powerful machine learning algorithm known for its effectiveness in regression and classification tasks. When applied to sales prediction, it can yield significant benefits for the company. By utilizing XGBoost, Big Mart can make data-driven decisions based on historical sales data, customer behavior, and other relevant information. This enables the

company to optimize various aspects of its operations

The below represents sales distributions

CONCLUSIONS

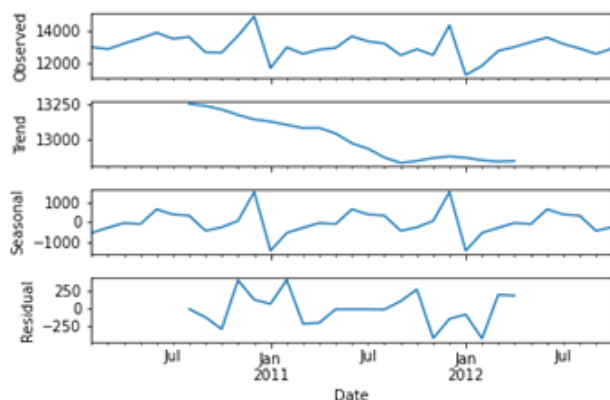
Nowadays, shopping malls and Big Marts maintain a comprehensive record of their sales data for each individual item. This data is utilized to predict future customer demand and update inventory management. These records are stored in data warehouses, containing a vast amount of customer data and item attributes. By mining this data, anomalies and frequent patterns can be identified. The resulting information can be leveraged to forecast future sales volume using various machine learning techniques, specifically for retailers like Big Mart. In this study, we propose a predictive model that utilizes the XG boost Regressor technique to forecast sales for companies similar to Big Mart. Our findings indicate that this model outperforms existing models in terms of performance.

The utilization of XGBoost for sales prediction in a retail setting like Big Mart is a valuable and data-driven approach. It empowers the company to make data-driven decisions, optimize operations, and enhance customer satisfaction. It's an essential tool for staying competitive and adapting to the dynamic retail landscape.

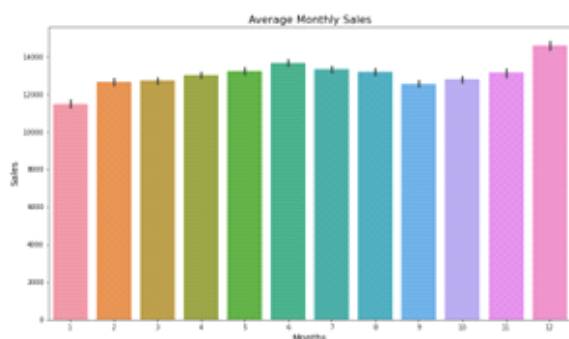
Sales predictions also aid in pricing strategies. Big Mart can set prices dynamically based on expected demand and inventory levels, potentially increasing profit margins.

TABLES AND FIGURES

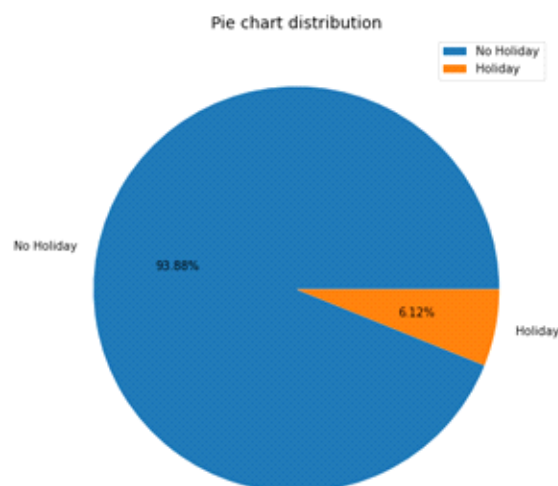
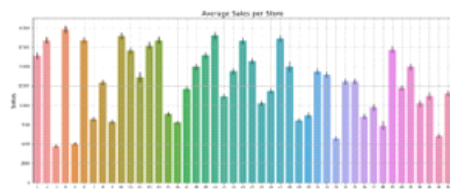
The below picture represents the seasonal sales Data



The below picture represent the monthly average sales data



The below picture represents the average sales in a store



REFERENCES

1. Smola, A., & Vishwanathan, S. V. N. (2008). Introduction to machine learning. Cambridge University, UK, 32, 34.
2. MacKay, D. J., & Mac Kay, D. J. (2003). Information theory, inference and learning algorithms. Cambridge university press.
3. Cerrada, M., & Aguilar, J. (2008). Reinforcement learning in system identification. In Reinforcement Learning. Intech Open.
4. Rohit Sav, Pratiksha Shinde, Saurabh Gaikwad: "Big Mart Sales Prediction Using Machine Learning "International Journal of Creative Research Thoughts (IJCRT)
5. Cheriyan, Sunitha, Shabina Ibrahim, Saju Mohanan, and Susan Treesa. "Intelligent Sales Prediction Using Machine Learning Techniques." In 2018 International Conference on Computing, Electronics&CommunicationsEngineering (icCECE), pp. 53-58. IEEE, 2018.
6. Quinlan, J. R. (2014). C4. 5: programs for machine learning. Elsevier.
7. T. Alexander and D. Christopher, "An Ensemble Based Predictive Modeling in Forecasting Sales of BigMart", International Journal of Scientific Research, vol. 5, no. 5, pp. 1- 4, 2016.